Generative Deep Learning for Targeted Compound Design

Abstract

In the past few years, de novo molecular design has increasingly been using generative models from the emergent field of Deep Learning, proposing novel compounds that are likely to possess desired properties or activities. De novo molecular design finds applications in different fields ranging from drug discovery and materials sciences to biotechnology. A panoply of deep generative models, including architectures as Recurrent Neural Networks, Autoencoders, and Generative Adversarial Networks, can be trained on existing data sets and provide for the generation of novel compounds. Typically, the new compounds follow the same underlying statistical distributions of properties exhibited on the training data set Additionally, different optimization strategies, including transfer learning, Bayesian optimization, reinforcement learning, and conditional generation, can direct the generation process toward desired aims, regarding their biological activities, synthesis processes or chemical features. Given the recent emergence of these technologies and their relevance, this work presents a systematic and critical review on deep generative models and related optimization methods for targeted compound design, and their applications.